Solubility prediction in mixed solvents: A combined molecular simulation and experimental approach

Solubility Prediction in Mixed Solvents: A Combined Molecular Simulation and Experimental Approach

Odin Kvam, Lev Sarkisov*

School of Engineering, University of Edinburgh, Edinburgh EH9 3FB, United Kingdom

Abstract

We propose a simple method to predict solubility in mixed solvents, combining experimental data for solubility in pure solvents with non-ideal contributions obtained from molecular simulation. By evoking a well-established thermodynamic model for mixed solvent, we provide rationale and justification for this hybrid approach. We test the accuracy of the method for solubility prediction in two highly non-ideal mixtures, CO_2 in ethanol + water, and CH_4 in MDEA + water. Non-ideal behaviour in each system is characterised using molecular simulation across the full range of compositions and for temperatures covering the range 273.15 K - 373.15 K. Comparison against experimental Henry's law constants for CO_2 and CH_4 gives mean absolute errors of 6.9% and 27%, respectively. We investigate the origin of non-ideal Henry's law behaviour in each of the mixed solvents by interrogating radial distribution functions from molecular simulation, finding increased association of CO_2 and CH_4 with the organic solvents at elevated temperatures.

Keywords: Solubility, Mixtures, Henry's Law, Molecular Simulation

1. Introduction

Solubility of gases in solvents is important for a wide range of chemical engineering processes, such as natural gas sweetening [1], waste water treatment [2], food processing [3], and post-combustion CO_2 capture [4]. Gas solubility in pure solvents is often readily available from the literature in the form of Henry's law constants, such as the compilation by Sander for water [5]. However, for mixed solvents, the substantial amount of solubility data accumulated over decades of industrial and academic research represents only a small subspace of possible combinations of mixture compositions, gases, and physical conditions. This motivates the development of predictive tools and theories for gas solubility in mixed solvents.

Classical molecular simulation offers direct prediction of thermodynamic, kinetic [6], and structural properties [7, 8] for mixed solvents of arbitrary composition. For prediction of gas solubility, molecular simulation has proven successful in both organic solvents and water. The quantitative accuracy of these predictions crucially depends on the quality of molecular models for solute and solvent species. Through calibration against experimental data, solute models can be optimised to

*Corresponding Author

Email address: lev.sarkisov@ed.ac.uk (Lev Sarkisov)

reproduce gas solubility in a particular solvent. By considering vapor-liquid equilibria of CO_2 and N_2 in alkanes, Potoff and Siepmann [9] proposed optimised gas models for use with the TraPPE-EH force field [10]. The same models were later applied to solubility in ethanol [11]. For amines and alkanolamines, Orozco et al. [12] have recently optimised solute parameters for physical solubility of CH_4, CO_2, and N_2O for use with the AUA force field [13]. In the case of water, Docherty et al. [14] detailed how optimised CH_4 parameters allows solvation thermodynamics to be correctly captured in TIP4P/2005 [15], a topic later explored in more detail by Ashbaugh et al. [16] by considering a range of water models. However, optimised solute models have limited transferability between solvents of different character, such as alkanes and water.

The issue of transferability is exacerbated for mixtures, where each pure solvent in principle requires separately optimised solute-solvent interactions. A similar approach may be used as for pure solvents by optimising each set of solvent-specific interaction parameters in turn. This was demonstrated by Orozco et al. [12] for CH_4 solubility, obtaining good agreement with the experimental Henry's law constant in 30% weight aqueous monoethanolamine at 303 K. However, ensuring a correct representation of solubility for each pure solvent amounts to a considerable effort in model development and validation, particularly for simulation of mixed solvents with more than two components.

We propose a hybrid approach for accurate solubility prediction in mixed solvents. It combines experimental data for Henry's law constants in pure solvents with molecular simulation in order to estimate non-ideal contributions to solubility in mixtures. Here, we explore the theoretical justification for such a hybrid approach, its advantages over direct molecular simulation, and potential limitations.

The starting point for our considerations is the variation of Henry's law constant with composition in mixed solvents. For a mixture with mole fractions x_i, Henry's law constant H can be described as [17]

$$\ln H = \sum_i x_i \ln H_i - \alpha, \tag{1}$$

where H_i corresponds to Henry's law constant in pure solvent i, and α is the deviation from an ideal mixture. The parameter α is often described in terms of the excess Henry's law constant H^E, with $\ln H^E = -\alpha$. In general, Henry's law constant is related to excess chemical potential μ_s^E for the solute by [18]

$$H = RT\rho_m \exp\left[\frac{\mu_s^E}{RT}\right], \tag{2}$$

where ρ_m is the molar density of the solvent, T is temperature, and R is the molar gas constant. Liquid density ρ is related to molar density by $\rho_m = \rho/M$, where average mixture molar mass M is obtained from pure solvent molar masses on mole fraction basis.

Conceptually, the solute excess chemical potential represents a balance between the strength of solute-solvent interactions and the cost of cavity creation in the solvent [19]. The variation of excess chemical potential with composition can be understood by considering the theory of O'Connell and Prausnitz [17] for a gas solute dissolving in a mixed solvent. In their theory, excess Gibbs energy G^E of the solute-solvent system is described as a function of mole fractions using a first order binary Margules expansion, of the form

$$\frac{G^E}{RT} = \sum_i \alpha_{si} x_s x_i + \sum_{i,j} \alpha_{ij} x_i x_j \tag{3}$$

where subscripts i, j indicate unlike solvents and s the gaseous solute. Coefficients α_{si} and α_{ij} encode ex-

cess free energies of binary solute-solvent and solvent-solvent interactions, respectively. The solute excess chemical potential is obtained by taking the derivative of excess Gibbs energy with respect to solute particles in the limit of vanishing concentration [17], giving

$$\mu_s^E = RT \sum_i \alpha_{si} x_i - RT \sum_{i,j} \alpha_{ij} x_i x_j \tag{4}$$

from which Henry's law constant can be calculated according to Eq. 2. In the case of a pure solvent, the last term in Eq. 4 vanishes. Hence, the meaning of each parameter α_{si} becomes clear: it is the reduced excess chemical potential $\alpha_{si} = \mu_s^E / RT$, encoding Henry's law constant in each pure solvent. Combining Eq. 2 and Eq. 4 gives

$$\ln H = \sum_i x_i \ln H_i - \sum_{i,j} \alpha_{ij} x_i x_j \tag{5}$$

where the second sum provides an explicit expression for α by comparison with Eq. 1. Hence, within the theory of O'Connell and Prausnitz [17] the parameter α is solute-independent, representing the reduced free energy of mixing for the solvent. Compared to an ideal mixture, solubility is expected to decrease in a solvent with negative free energy of mixing, and increase for a solvent with positive free energy of mixing. This suggests that solute-solvent interactions play a limited role for solubility in mixed solvents, with this property primarily governed by contributions from pure solvents.

As emphasised by O'Connell and Prausnitz [17], the symmetric Margules expansion in Eq. 3 is only a good assumption for mixtures of simple non-polar solvents. Nonetheless, mixture excess energies appear to be reasonably symmetric functions, both for aqueous ethanol [20] and in a recent simulation study of aqueous alkanolamines [21]. This suggests a degree of transferability to generic mixtures. Indeed, Eq. 3 is not the only thermodynamic model resulting in a solute-independent expression for α, with similar approaches to estimating solubility in mixed solvents from non-ideal behaviour investigated by Kung et al. [22] based on variations of the Wilson equation [23], and Shulgin and Ruckenstein [24] using Kirkwood-Buff integrals. In each case solubility in the mixed solvent is expressed as a variant of Eq. 1, with α represented by different functional forms. In practice, the solute-independent character of α extends only to molecules of similar size and shape, due to Eq. 3 ignoring ternary and higher terms for solute-solvent interactions.

The separation of contributions to solubility in mixed solvents is at the heart of our approach. However, we do not attempt to correlate solubility using Eq. 5, due to the inherent limitations of the theory explored above. Instead, we directly employ Eq. 1 for solubility prediction, using molecular simulation to obtain α at each condition of interest. Simulated values of α are combined with pure solvent Henry's law constants H_i from experimental sources, leveraging the considerable literature on gas solubility in pure solvents. In essence, this process eliminates the need for solute-solvent parameter optimisation by separating contributions of mixture thermodynamics from solubility in pure solvents.

In the following sections, we explore the accuracy of Eq. 1 for solubility prediction in aqueous-organic solvents on the basis of α obtained by molecular simulation. We consider two cases. First, solubility of CO_2 in aqueous ethanol is compared against the experimental results of Dalmolin et al. [25] and Postigo and Katz [26], each covering the full range of compositions at several temperatures. Second, solubility of CH_4 in aqueous methyldiethanolamine (MDEA) is compared against the experimental results of Jou et al. [27], Jou

et al. [28], and Schmidt et al. [29] for pure MDEA and mixtures with 34.7% and 50% MDEA by weight, respectively. Further, we investigate the degree to which α is a solute-independent property in aqueous ethanol by comparison between N_2 and CO_2, two small linear molecules with vastly different solubility in each of the pure solvents ethanol and water.

2. Methods

2.1. Force Fields and Molecular Systems

Force field parameters for solvents and gases are selected from the existing literature on simulation of liquid phase systems and vapour-liquid equilibria. For solvents, we prioritise a correct representation of liquid densities. This is motivated by the sensitivity of solute excess chemical potentials to solvent density, an issue investigated by Garde et al. [30] and later explored by Ashbaugh et al. [16] for CH_4 solubility in water. We avoid models specifically optimised for the mixed solvents of interest, such as the AUA force field of Orozco et al. [13] for alkanolamines and their mixtures: our objective is to accurately predict mixture Henry's law constants informed only by experimental data for the behaviour of pure solvents. In general, stronger predictive capabilities should be expected from models optimised for specific systems.

For organic solvents, a large number of molecular models are available. Since we wish to consider generic liquid mixtures for which there may not be optimised models available, we use the well-established OPLS/AA force field [31, 32], selected based its wide coverage of chemical space and a systematic comparison against experimental isothermal-isobaric properties of pure liquids [33]. OPLS/AA shows particularly good agreement for liquid densities at ambient conditions, with an average deviation of 2% across 235 organic liquids. For CH_4, N_2, and CO_2, we use TraPPE [9, 34] Lennard-Jones parameters, charges, and equilibrium geometries, with a united-atom representation of CH_4. We additionally use flexible bonded parameters for CO_2 [35] and N_2, detailed in Supporting Information. Water is represented by TIP4P/2005 [15], found to accurately reproduce a number of isothermal-isobaric liquid properties [36] and successfully used to capture the temperature dependence of CH_4 solubility [16, 14]. The rigid geometry of TIP4P/2005 is maintained with the analytical SETTLE [37] algorithm. Geometric combination rules are employed for Lennard-Jones parameters between unlike atom types. For ethanol and MDEA, 1-4 intramolecular Lennard-Jones and coulombic interactions are scaled by a factor of 0.5 in accordance with the OPLS/AA force field. To limit the variation of simulated properties with system size [38], total mass is maintained close to that of 800 water molecules.

Solvent properties and Henry's law constants are probed over a grid of compositions and temperatures both for ethanol + water mixtures and MDEA + water mixtures. In addition to the special case of pure solvents, each mixture is simulated at 0.1 weight fraction increments, giving 11 distinct system compositions. 16 geometrically spaced temperatures are used to cover the range 273.15 K - 373.15 K, producing a total of 176 temperature-composition grid points for each mixture. For comparison against experimental data at specific temperatures and compositions, we linearly interpolate between simulated grid points using Delaunay triangulation [39].

2.2. Simulation of Liquid Properties

Our choice of methodology for obtaining liquid phase properties of solvents is motivated by two requirements. Firstly, the solvent components must be allowed to sufficiently mix to reflect true equilibrium systems. Secondly, densities must be determined over a wide range of temperatures. These requirements naturally suggest replica exchange molecular dynamics [40–43] (REMD) as the method of choice. The use of REMD with replica systems in a ladder of increasing temperatures accelerates convergence towards equilibrium states by leveraging rapid exploration of configurational space at high temperatures to overcome energy barriers at low temperatures. For each mixture composition, we simulate the temperatures of interest with 16 independently generated systems, attempting exchange of coordinates and velocities between systems of adjacent temperatures at 1 ps intervals. Particle velocities are scaled after each successful exchange to maintain correct Maxwell-Boltzmann distributions for kinetic energy. Resulting exchange probabilities are in the range 20% - 30% for all systems considered. A total of 20 ns is allowed for REMD simulation, with the final 10 ns used for data collection. Uncertainty estimates for liquid properties are obtained from 2 ns block averages.

All systems are simulated under periodic boundary conditions in the isobaric-isothermal ensemble. Equations of motion are integrated with a velocity-Verlet algorithm [44] at a time step of 1 fs. Isothermal-isobaric conditions are maintained using the Nosé-Hoover [45, 46] thermostat with coupling time 2 ps and the Martyna-Tobias-Tuckerman-Klein [47] barostat, selected for compatibility with the velocity-Verlet integrator, with coupling time 5 ps. Coulomb potentials are truncated at 0.9 nm with a smooth particle-mesh Ewald method [48, 49] employed for long range interactions, using a grid spacing of 0.1 nm and cubic interpolation. Lennard-Jones potentials are similarly truncated at 0.9 nm, with analytical tail corrections [50] added for system energy and pressure. Simulations are carried out using GROMACS [51] version 5.0.7.

2.3. Simulation of Excess Chemical Potentials

A range of methods have been successfully applied to simulate gas solubility in molecular systems, including Gibbs ensemble Monte Carlo [52, 11], the method of expanded ensembles [53], and a variety of free energy perturbation approaches of which the Widom particle insertion method [54, 18] represents the most straight forward example. In each case, solubility is obtained by considering transitions of the solute between the gas and liquid phases. Depending on the system, transitions may be attempted directly, using configurational bias rules [55], or through intermediate thermodynamic states with increasing solute-solvent coupling.

In this work, we employ the multistate Bennet acceptance ratio [56] (MBAR) perturbation method in combination with expanded ensemble sampling of solute-solvent interactions. We use a chain of finely spaced thermodynamic states, dynamically explored over the course of each simulation. This approach has several advantages for solubility estimation in dense heterogeneous liquids. Firstly, variable solute-solvent coupling enhances solute mobility, avoiding repeated sampling of the same solvent environment. Secondly, the method is insensitive to solute size, avoiding the convergence concerns of e.g. the Widom method in dense systems. Finally, it permits the use of configurations obtained from simulation of isothermal-isobaric solvent properties as starting points, since the expanded ensemble description

naturally includes thermodynamic states with no solute-solvent coupling.

Interaction potentials between the solute and solvent are gradually switched on, defining a pathway of intermediate thermodynamic states. For a sufficiently small perturbation in interaction potentials $\Delta U = U_n(\mathbf{r}) - U_m(\mathbf{r})$ from state m to n at system coordinates $\mathbf{r}$, the corresponding free energy change $\Delta_{m\to n}G$ can be calculated as

$$\Delta_{m\to n}G = RT \ln \left\langle \exp\left[\frac{\Delta U}{RT}\right]\right\rangle_m , \qquad (6)$$

where the angled brackets indicate an average over the configurational ensemble of state m. Potential energy differences are calculated using scaled potentials: Lennard-Jones potentials are scaled using a 1-1-6 potential [57] with soft core parameter [58] 0.50, and Coulomb interactions are scaled using a linearly interpolated standard Coulomb potential, with further detail in Supporting Information. Lennard-Jones and Coulomb interactions are coupled orthogonally to avoid the large forces associated with closely spaced charges [59]. We use uniformly spaced scaling parameters at intervals of 0.02 for Lennard-Jones potentials and 0.05 for Coulomb potentials, yielding 51 thermodynamic states for the neutral CH_4 molecule, and 71 states for CO_2 and N_2.

Potential energy differences between all thermodynamic states are sampled in a single simulation using the method of expanded ensembles [53, 60], allowing the solute to couple or decouple using Monte Carlo transitions between thermodynamic states. We use Gibbs sampling [61] of thermodynamic states with transitions attempted every 0.1 ps, and accepted according to the Metropolis criterion

$$P(m \to n) = \min \begin{cases} \exp\left[-(U_n^* - U_m^*)/RT\right] \\ 1 \end{cases} , \qquad (7)$$

where $U_m^* = U_m(\mathbf{r}) + \eta_m$ is defined by the instantaneous potential energy $U_m(\mathbf{r})$ and expanded ensemble weight η_m of state m. For even exploration of all states, weights should be chosen such that $\eta_n - \eta_m = \Delta_{m\to n}G$. Since this free energy difference is not known *a priori*, we find approximate values for η_m during a 2 ns calibration period by cumulatively averaging the energy difference in Eq. 7 with weights initially assigned as zero for all states. Data collection is carried out over a further 15 ns of simulation, with potential energies for all thermodynamic states saved every 0.1 ps. The potential energy differences obtained from expanded ensemble sampling are analysed using MBAR [56] to calculate $\Delta_{m\to n}G$ for each pair of thermodynamic states. By considering the free energy change for the complete transition from gas to solvent, μ_s^E may be obtained for the solute.

2.4. Calculation of Henry's Law Constants

Having obtained ρ and μ_s^E, Eq. 2 can be used directly to calculate Henry's law constant in the solvent. However, in order to investigate the variation of Henry's law constant and variation of α as a smooth function of temperature and compositions, we employ smooth cubic bivariate B-spline fits to ρ and μ_s^E. Further, the statistical variance of the predicted Henry's law constants is analysed using the technique of bootstrap resampling [62]. For each solute-solvent combination, the set of 176 independent simulation results is repeatedly bootstrap resampled and fitted to a cubic bivariate B-spline as done for ρ. By considering 10,000 bootstrapped spline fits of the simulation data, a maximum likelihood value and

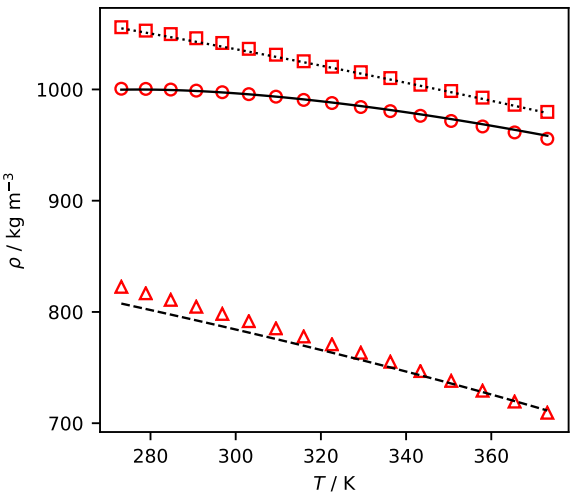

Figure 1: Liquid mass density as a function of temperature for pure solvents. Open symbols correspond to simulated values; ○: water, △: ethanol, and □: MDEA. Lines correspond to empirical correlations; — water [63], – – ethanol [64], and ··· MDEA [65].

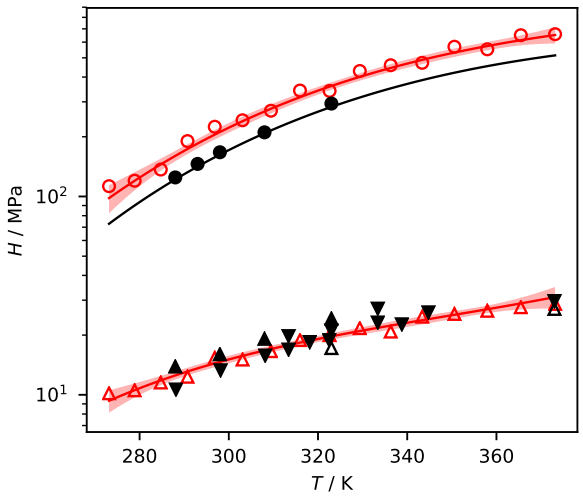

Figure 2: Henry's law constant H as a function of temperature for CO_2 in pure solvents. Open symbols correspond to simulated values; ○: water, △: ethanol, △: TraPPE CO_2 in TraPPE-UA ethanol, reported by Zhang et al. [11]. Filled symbols correspond to literature experimental values; ●: water [25], ▲: ethanol [25], ▼: ethanol [66–69]. — Semiempirical correlation of Harvey [70] for water. — Maximum likelihood estimate, with the shaded area representing 95% confidence intervals.

95% confidence interval is calculated for each point of temperature and composition, assuming the fitted values of μ_s^E follow a Gaussian distribution. Examples of the cubic bivariate B-spline fit and bootstrapping method used to calculate maximum likelihood values and confidence intervals are provided in Supporting Information.

3. Results and Discussion

3.1. Pure Solvents

To assess the non-ideal contributions to gas solubility in mixtures, we first consider properties of the pure solvents. This section reports simulated densities for water, ethanol, and MDEA, together with Henry's law constants calculated using Eq. 2 for CO_2, N_2, and CH_4 in pure solvents. In all figures, red open symbols indicate our simulated values, black open symbols indicate simulated values from other authors, and filled symbols indicate experimental values.

Pure solvent liquid densities are shown in Fig. 1, with simulated results from this work compared against experimental correlations for water [63], ethanol [64], and MDEA [65]. Simulated results for water and MDEA show mean absolute deviations from experimental values of 0.10% and 0.16%, respectively. Ethanol shows a larger deviation of 1.0%, with a maximum deviation of 1.9% at 273.15 K. Simulated values for water density are found to be in good agreement with the results of Abascal and Vega [15].

Simulated Henry's law constants for CO_2 in ethanol and water are compared against available experimental values in Fig. 2. In ethanol, Dalmolin et al. [25] report an increase in Henry's law constant from 10 MPa to 20 MPa over the temperature range 288 K - 323 K. This trend is corroborated by other experimental

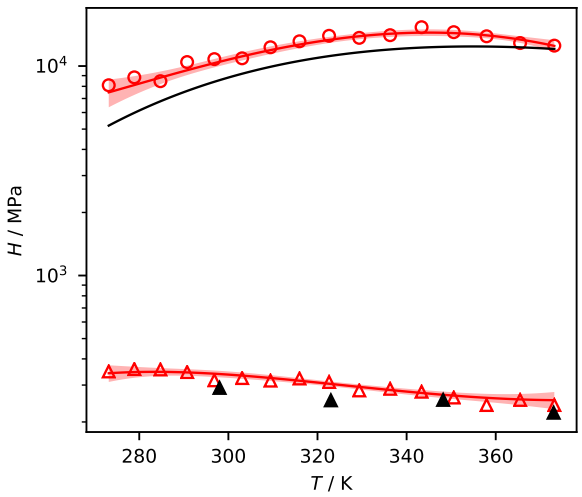

Figure 3: Henry's law constant H as a function of temperature for N_2 in pure solvents. Open symbols correspond to simulated values; ○: water, △: ethanol. Filled symbols correspond to literature experimental values; ▲: ethanol [66, 68]. — Semiempirical correlation of Harvey [70] for water. — Maximum likelihood estimate, with the shaded area representing 95% confidence intervals.

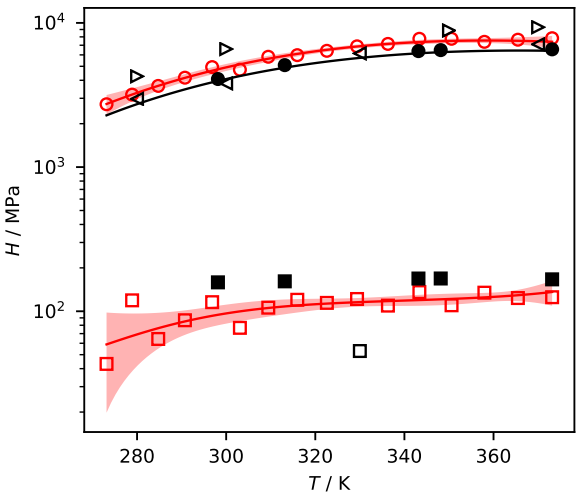

Figure 4: Henry's law constant H as a function of temperature for CH_4 in pure solvents. Open symbols correspond to simulated values; ○: water, □: MDEA, ▷: water, reported by Docherty et al. [14], ◁: optimised CH_4 parameters for solvation in water [14], □: AUA model of MDEA, reported by Orozco et al. [12]. Filled symbols correspond to literature experimental values; ●: water [72], ■: MDEA [27]. — Semiempirical correlation of Harvey [70] for water. — Maximum likelihood estimate, with the shaded area representing 95% confidence intervals.

sources [66–69], although some scattering is apparent. Simulated Henry's law constants in ethanol are in good agreement with experimental data, with a similar agreement previously obtained by Zhang and Siepmann [11] for TraPPE CO_2 using a united-atom representation of ethanol. In water, the experimental results of Dalmolin et al. [25] are in good agreement with the semiempirical correlation of Harvey [70]. Simulated Henry's law constants are in this case systematically offset from experimental values, with a a deviation on the order of 50%.

It is worth noting that our simulations of CO_2 in aqueous systems do not attempt to account for the formation of carbonic acid derivatives through hydration reactions. Solubility is calculated on the basis of molecular CO_2, in the same manner as for the inert solutes CH_4 and N_2. For experimental measurements of Henry's law constants in neutral systems, the equilibrium reactions of CO_2 can safely be ignored [71]. Hence, direct comparison between experimental sources and results from molecular simulation appears justified for both ethanol and water, as well as their mixtures.

Simulated Henry's law constants for N_2 in ethanol and water are shown in Fig. 3, together with reported experimental values for N_2 in ethanol [66, 68]. Results from simulation show a reduction of Henry's law constant with temperature in ethanol, consistent with the experimental data. Our simulated values for Henry's law constant in water show a similar offset from the semiempirical correlation of Harvey [70] as seen for CO_2. For both gas species, solute-solvent interactions are underestimated in water.

In Fig. 4, simulated Henry's law constants for CH_4 in pure solvents are compared against the experimental results of Jou et al. [27] for MDEA and the semiempirical correlation of Harvey [70] for water. Solvation of CH_4 in water is strongly disfavoured due to its nonpolar character, with a Henry's constant increasing from 2.3 GPa to 6.4 GPa over the temperature range. As seen for both CO_2 and N_2, simulated Henry's law constants for CH_4 in water are consistently overestimated compared to experimental values.

In this case, we find an offset on the order of 20% compared to experimental Henry's law constants. This result is somewhat lower than the values reported by Docherty et al. [14] in their simulation using TIP4P/2005 and the Hirschfelder model of CH_4, included in Fig. 4 for comparison. Also shown are Henry's law constants for a CH_4 model optimised to reproduce experimental solubility in TIP4P/2005 [14]. In both comparisons, Henry's law constants are calculated from the reported simulation data using Eq. 2. For validation purposes we reproduce the result of Docherty et al. [14] at 280 K, obtaining a result of 3.9 ± 0.5 GPa, in satisfactory agreement with the reported value of 4.3 GPa.

Solubility of CH_4 in MDEA is considerably greater than in water. Jou and Mather [27] report Henry's law constants for CH_4 in MDEA near 165 MPa for temperatures between 298.15 K and 373.15 K. We find simulated Henry's law constants to be somewhat lower, with values near 100 MPa for temperatures above 300 K. No clear trend can be deduced below 300 K. Our simulated results indicate a significantly higher Henry's law constant at 330 K than the value of 54 ± 4 MPa reported by Orozco et al. [12] using the AUA force field, included in Fig. 4 for comparison. Contrary to the result for water, Henry's law constant is underestimated by approximately 50%. The simultaneous over- and underprediction of Henry's law constant for CH_4 in respectively water and MDEA illustrate the lack of transferability for solute models between chemically different solvents: increasing solute-solvent interactions to reach agreement for solubility in water will worsen predictions for MDEA, necessitating solvent-specific interaction parameters.

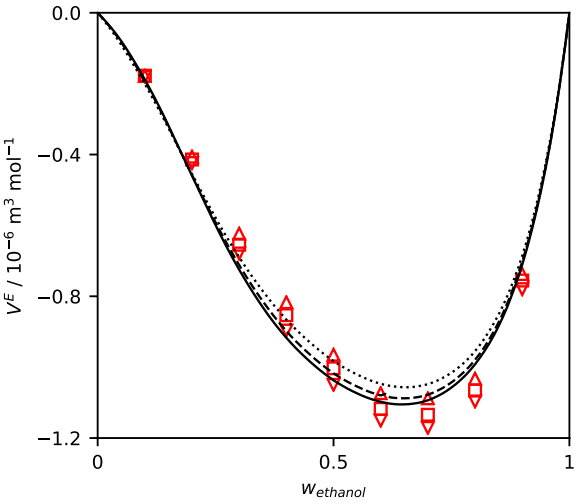

Figure 5: Excess molar volume V^E as a function of composition for ethanol + water mixtures. Open symbols correspond to simulated values; ▽: 288.15 K, □: 298.15 K, and △: 308.15 K. Lines correspond to literature experimental values [64]; — 288.15 K, – – 298.15 K, and ··· 308.15 K.

3.2. *CO_2 in Ethanol + Water*

Excess molar volumes for ethanol + water mixtures are shown as a function of composition in Fig. 5, with our simulated results compared against experimental values [64] at temperatures of 288.15 K, 298.15 K, and 308.15 K. Simulated values have a mean uncertainty of 9×10^{-9} m^3 mol^{-1}. For all solvent mixtures, composition is indicated by weight fraction w of the organic solvent. Simulated excess molar volumes are in near-

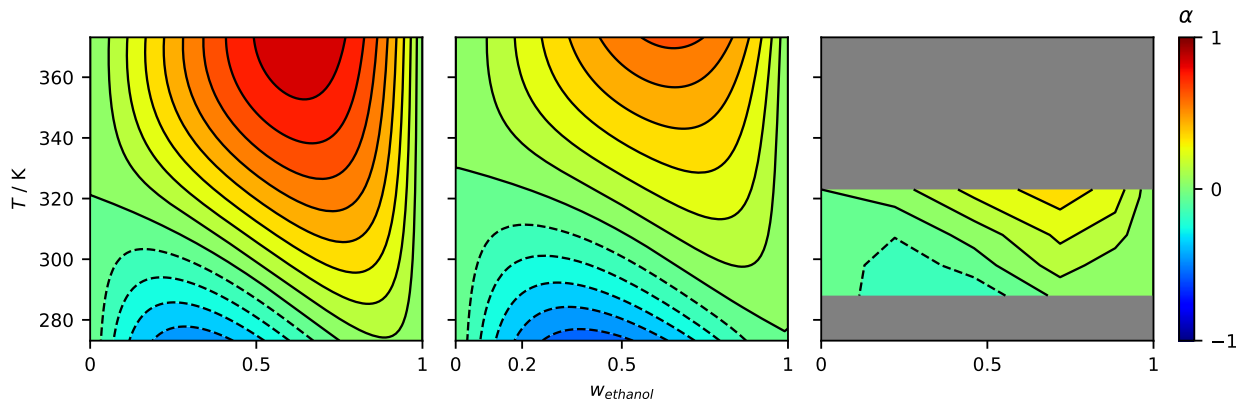

Figure 6: Non-ideality parameter α in ethanol + water systems, calculated from Henry's law constants on the basis of Eq. 1. Left panel: simulated data for N_2. Centre panel: simulated data for CO_2. Right panel: experimental data for CO_2 from Dalmolin et al. [25]. Contours are plotted at increments of 0.1, with dashed contour lines indicating negative α values.

quantitative agreement with the real mixture, suggesting the variation of Henry's law constant in the mixture will be accurately captured. Hence, we anticipate good agreement for the non-ideality parameter α between molecular simulation and experimental data. However, as shown in Supporting Information the mixture excess enthalpy is significantly underestimated for water + ethanol systems.

In Fig. 6, the non-ideal contribution to Henry's law constant, α, is shown as a function of temperature and composition for ethanol + water mixtures. The left panel shows values of α obtained by simulation of N_2, while the centre panel shows values obtained by simulation of CO_2. Despite a qualitatively different behaviour for Henry's law constant in ethanol and a difference of nearly two orders of magnitude in solubility seen between Fig. 2 and Fig. 3, values of α for N_2 closely resemble those for CO_2 in this non-ideal system. For both solutes, α is calculated using cubic bivariate B-spline fits of ρ_m and μ_s^E, detailed in Supporting Information. From the root mean squared errors of the fit we obtain uncertainties of 0.6 kg m^{-3} for density, 0.14 kJ mol^{-1} for N_2 excess chemical potential, and 0.16 kJ mol^{-1} for CO_2 excess chemical potential.

In the right panel of Fig. 6, the parameter α calculated from the experimental data of Dalmolin et al. [25] is shown for comparison. We observe good correspondence between simulated and experimental values for CO_2, and the same features appearing in all three panels. At low temperatures and ethanol content, there is a region of negative non-ideal behaviour, corresponding to a reduction in solubility compared to the ideal mixture behaviour ($\alpha = 0$). As temperature increases there is a transition to positive values of α, corresponding to enhanced solubility compared to the ideal mixture. The temperature of this transition is clearly dependent on concentration. This illustrates the deficiency of the Margules expansion for describing real mixed solvents: the functional form of Eq. 3 and a a fixed binary interaction parameter α_{ij} implies that the behaviour of a mixture is symmetric for all concentrations at a given temperature. Conversely, α defined in Eq. 1 and obtained by molecular simulation is not constrained by any functional expression.

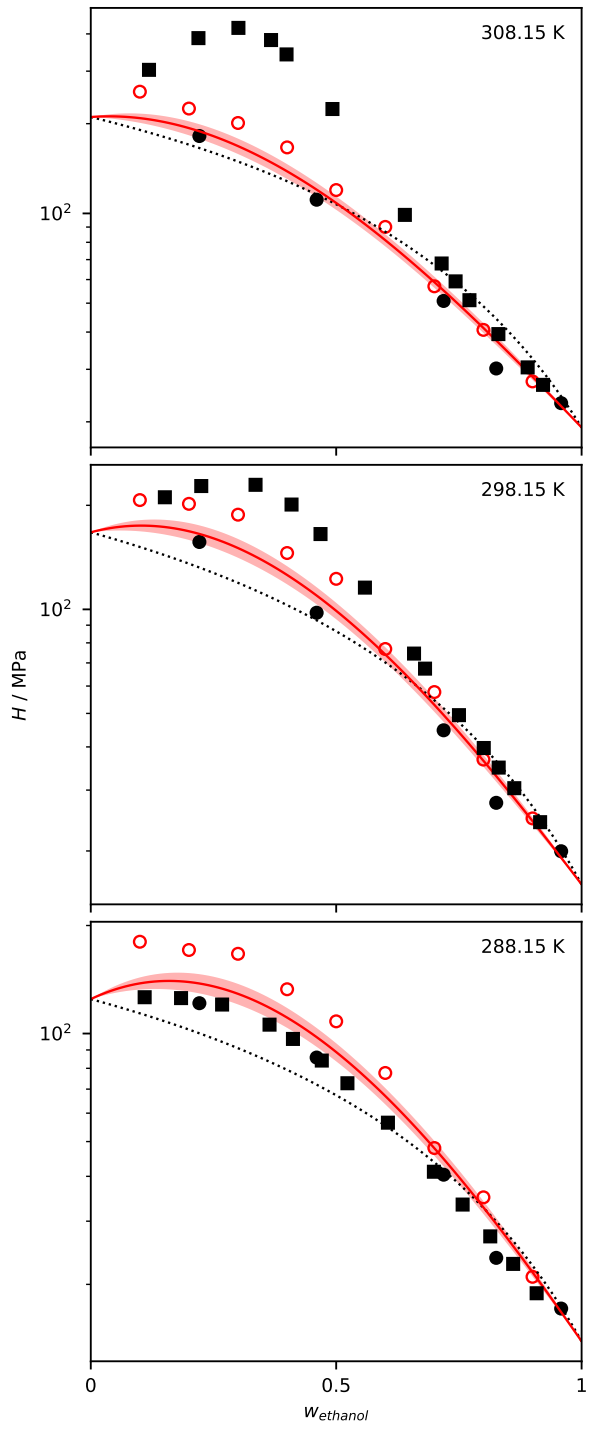

Figure 7: Henry's law constant H as a function of composition for CO_2 in ethanol + water mixtures at 288.15 K, 308.15 K, and 323.15 K. ○: Simulated values, ●: experimental values from Dalmolin et al. [25], ■: experimental values from Postigo and Katz [26], ··· ideal mixture, — prediction on basis of Eq. 1, using α obtained from molecular simulation and Henry's law constants for pure solvents from Dalmolin et al. [25]. The shaded area represents 95% confidence intervals.

Based on the values of α calculated from molecular simulation, Henry's law constants for the mixture may be predicted using Eq. 1 in combination with experimental pure solvent data. Fig. 7 shows predictions using this hybrid approach for CO_2 in ethanol + water mixtures at temperatures of 288.15 K, 298.15 K, and 308.15 K, in comparison with experimental data from Dalmolin et al. [25] and Postigo and Katz [26]. Dotted lines indicate ideal mixture behaviour based on pure solvent Henry's law constants [25]. The experimental sources are in close agreement at 288.15 K, showing a considerable increase in Henry's law constant compared to the ideal mixture for concentrations below 0.5 ethanol weight fraction. This behaviour corresponds to the area of negative α values in Fig. 6. At temperatures of 298.15 K and 308.15 K there is a qualitative difference between the experimental sources, despite good agreement for CO_2 solubility in both pure solvents. Postigo and Katz [26] find increasingly non-ideal behaviour with increasing temperatures, while Dalmolin et al. [25] report Henry's law constants closer to the ideal mixture.

Simulated Henry's law constants in water are overpredicted at each temperature, as seen previously in Fig. 2. For mixed solvents, the offset observed for water is carried over in proportion to concentration. Simulated Henry's law constants are higher than both experimental sources at 288.15 K, and fall between the two sets of data at 298.15 K and 308.15 K. We obtain a mean absolute deviation of 19% compared against both pure and mixed solvent data, with deviations in excess of 30% for low ethanol weight fraction mixtures.

Henry's law constants obtained using Eq. 1 with pure solvent data from Dalmolin et al. [25] and simulated values of α are shown as red dashed lines in Fig. 7.

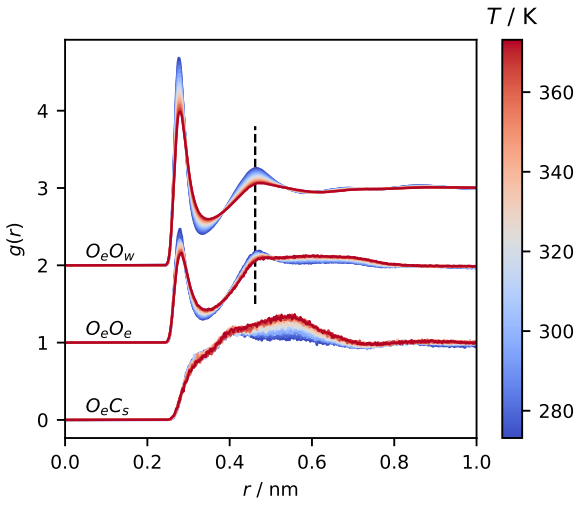

Figure 8: Variation of radial distribution functions with temperature in ethanol + water at 0.2 ethanol weight fraction. The value of $g(r)$ for O_eO_e and O_eO_w is shifted by 1 and 2 respectively. Dashed line connecting secondary peaks in $g(r)$ for O_eO_w and O_eO_e to guide the eye.

By tethering pure solvent values to experimental data, mixed solvent predictions are brought in line for the whole composition range. We obtain a mean absolute percentage error of 6.9% for prediction of mixed solvent Henry's law constants compared against the results of Dalmolin et al., [25] similar to the observed scattering in experimental sources for ethanol in Fig. 2.

As previously noted, non-ideal contributions to solubility in mixed solvents predominantly originates from solvent-solvent interactions, with solute-solvent interactions playing a minor role. By considering the thermodynamics of ethanol + water mixtures, we may gain some insight into the temperature- and composition-dependent variation of α observed in Fig. 6. In previous studies, an unusual liquid structure has been reported for aqueous alcohol mixtures at low temperatures [73–75], coinciding with a minimum in excess entropy of the mixed solvent at low ethanol weight fraction [20]. Upon mixing, the two solvents form intercalated clusters, increasing order in the structure of the liquid. This increase in structure was recently investigated by Lenton et al. [75], concluding that molecular segregation is driven by hydrogen bond formation and the isolation of hydrophobic groups. If these structural features are the source of negative non-ideal behaviour in Fig. 6, we expect to see a change in liquid structure as α transitions to positive values with increasing temperature.

To investigate the liquid structure of ethanol + water, it is instructive to consider the radial distribution function $g(r)$ [50] for selected atom pairs in ethanol + water mixtures. The radial distribution function is a measure of association, showing relative concentration as a function of separation. Fig. 8 illustrates the variation of $g(r)$ across the temperature range 273.15 K - 373.15 K for selected atom pairs at 0.2 ethanol weight fraction. Due to the importance of polar group interactions [75] we examine $g(r)$ for the atom pairs O_eO_e and O_eO_w, where capital letters indicate element and subscripts e and w indicate ethanol and water, respectively. At low temperatures, both O_eO_e and O_eO_w radial distribution functions show clear primary peaks near 0.25 nm and secondary peaks near 0.46 nm, indicating the two solvent species form a shared liquid structure [75]. As temperature increases the secondary peak for O_eO_e recedes. We observe a transition to more loosely organised ethanol clusters above 325 K, evidenced by the plateau in $g(r)$ at 0.5 nm - 0.7 nm. This temperature coincides with α inverting from negative to positive values for 0.2 weight fraction ethanol in the centre panel of Fig. 6.

We additionally consider the association of CO_2 with ethanol, due to the large difference in solubility between the two pure solvents. A separate set of isothermal-

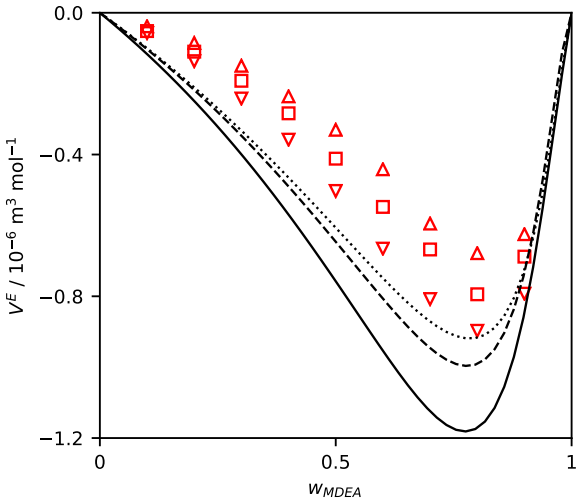

Figure 9: Excess molar volume V^E as a function of composition for MDEA + water mixtures. Open symbols correspond to simulated values at temperatures of: ▽: 298.15 K, □: 343.15 K, and △: 373 K. Lines correspond to the experimental correlation of Bernal-García et al. [65] at temperatures of: — 298.15 K, – – 343.15 K, and ··· 373.15 K (extrapolated).

isobaric simulations are used to obtain configurations for the solute-solvent system, since only a small fraction configurations from expanded ensemble simulations correspond to fully coupled solutes. Fig. 8 shows $g(r)$ for the atom pair O_eC_s, where subscript s indicates the solute species. As temperature increases CO_2 associates more strongly with ethanol, with a peak in $g(r)$ near 0.55 nm at 373.15 K. It appears that the enhanced solubility of CO_2 at elevated temperatures is due to preferential association with ethanol clusters.

3.3. CH_4 in MDEA + Water

Excess molar volumes for MDEA + water mixtures are plotted as a function of composition in Fig. 9, with our simulated values compared against the experimental results and correlation of Bernal-García et al. [65] at 298.15 K, 343.15 K, and 373.15 K. Simulated values have a mean uncertainty of 1.3×10^{-8} m^3 mol^{-1}. The correlation at 373.15 K is calculated by extrapolation of MDEA density and mixture excess molar volumes from the reported range of 283.15 K - 363.15 K. Despite accurate representation of both pure solvent densities in Fig. 1, molecular simulation systematically predicts excess molar volumes that are overly positive for these mixtures, reflecting the limited accuracy of solvent-solvent interactions. Consequently, we expect non-ideal contributions to mixture solubility to be overly negative in MDEA + water systems. This contrasts with the results of Orozco et al. [76] using the AUA force field in combination with TIP4P/2005, showing excess molar volumes that are overly negative with a minimum near -2.5×10^{-6} m^3 mol^{-1} at 298 K. However, as shown in Supporting Information, mixture excess enthalpy appears to be accurately described in our simulations.

As for ethanol + water, the non-ideality parameter α calculated for CH_4 solubility in MDEA + water is shown in Fig. 10 as a function of temperature and composition. The left panel shows values obtained from molecular simulation. In this case, only positive non-ideal behaviour is observed, indicating increased mixture solubility over the ideal case for all compositions and temperatures. Values are calculated using the same cubic bivariate B-spline method as for ethanol + water systems, with root mean squared errors of 0.3 kg m^{-3} and 0.32 kJ mol^{-1} for ρ and μ_s^E, respectively.

The non-ideal contribution to solubility of CH_4 in MDEA + water systems has previously been investigated by Schmidt et al. [29] based on experimental Henry's law constants in pure solvents [27, 72] together with mixtures with weight fractions of 0.347 MDEA [28] and 0.5 MDEA [29]. Corresponding experimental values of α are shown in the centre panel of Fig. 10 for temperatures up to 373.15 K. While

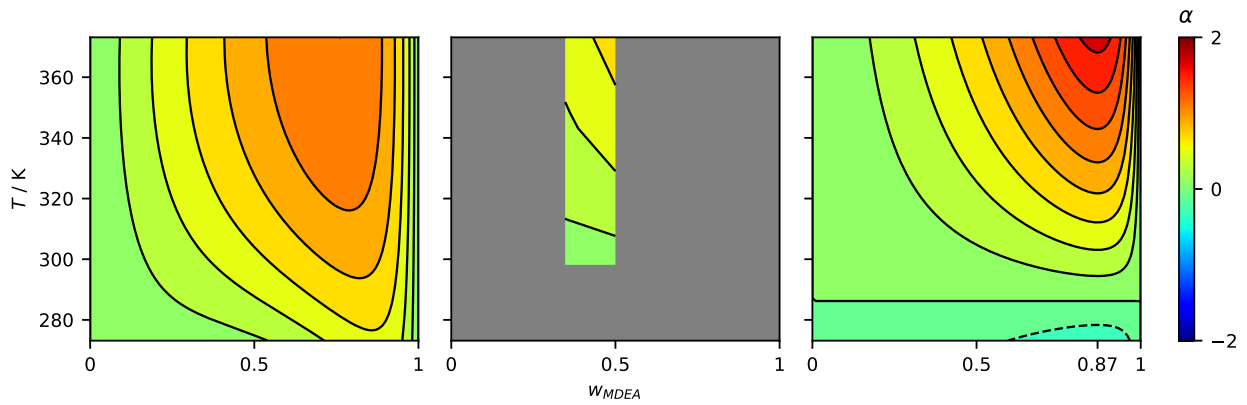

Figure 10: Non-ideality parameter α in MDEA + water systems, calculated from Henry's law constants on the basis of Eq. 1. Left panel: simulated data for CH_4. Centre panel: experimental data for CH_4 [29, 28, 27, 70]. Right panel: temperature correlation of Schmidt et al. [29] applied to the full composition range. Contours are plotted at increments of 0.2, with dashed contour lines indicating negative α values.

there is limited data for comparison, values for α appear lower than those obtained from molecular simulation. Schmidt et al. [29] correlated mixture solubility using a temperature dependent Margules parameter α_{ij}, as defined in Eq. 5. The right panel of Fig. 10 shows α calculated from their correlation. In this case, mixture non-ideality sharply peaks at 0.87 MDEA weight fraction, outside the range of available experimental data. The peak corresponds to $x_{MDEA} = 0.5$, and is a consequence of the symmetric representation of mixture excess Gibbs energy in Eq. 3. Additionally, there is an inversion of α from positive to negative values as temperature reduces below 286.22 K, not seen in molecular simulation. As the data used for the correlation only extends downwards to 298.15 K it is uncertain whether this inversion would be observed experimentally.

Predicted Henry's law constants of CH_4 in MDEA + water are shown in Fig. 11 at temperatures of 298.15 K, 343.15 K, and 373.15 K, calculated using Eq. 1 with α from molecular simulation and experimental pure solvent data [27, 70]. Predictions are compared against reported experimental Henry's law constants at 34.7% [28] and 50% [29] MDEA weight fraction. As temperature increases, both experimental and simulated Henry's law constants increasingly deviate from the ideal mixture, illustrated by the dotted lines based on experimental pure solvent data [27, 72]. However, simulations systematically underestimates mixture Henry's law constants at each temperature. We obtain a mean absolute percentage deviation of 22% compared against both pure and mixed solvent data.

Further, Henry's law constants obtained using Eq. 1 with experimental pure solvent data [27, 72] and α obtained from molecular simulation are indicated by dashed red lines in Fig. 11. In this case comparison with experimental mixed solvent gives a mean absolute error of 27%, worse than direct prediction from molecular simulation. This result is not surprising considering the under-prediction of excess molar volumes in Fig. 9: incorrectly represented mixture thermodynamics leads to poor predictions for solubility, regardless of the quality of solute-solvent interaction parameters.

The black dashed lines in Fig. 11 indicate Henry's law constants calculated from the temperature-

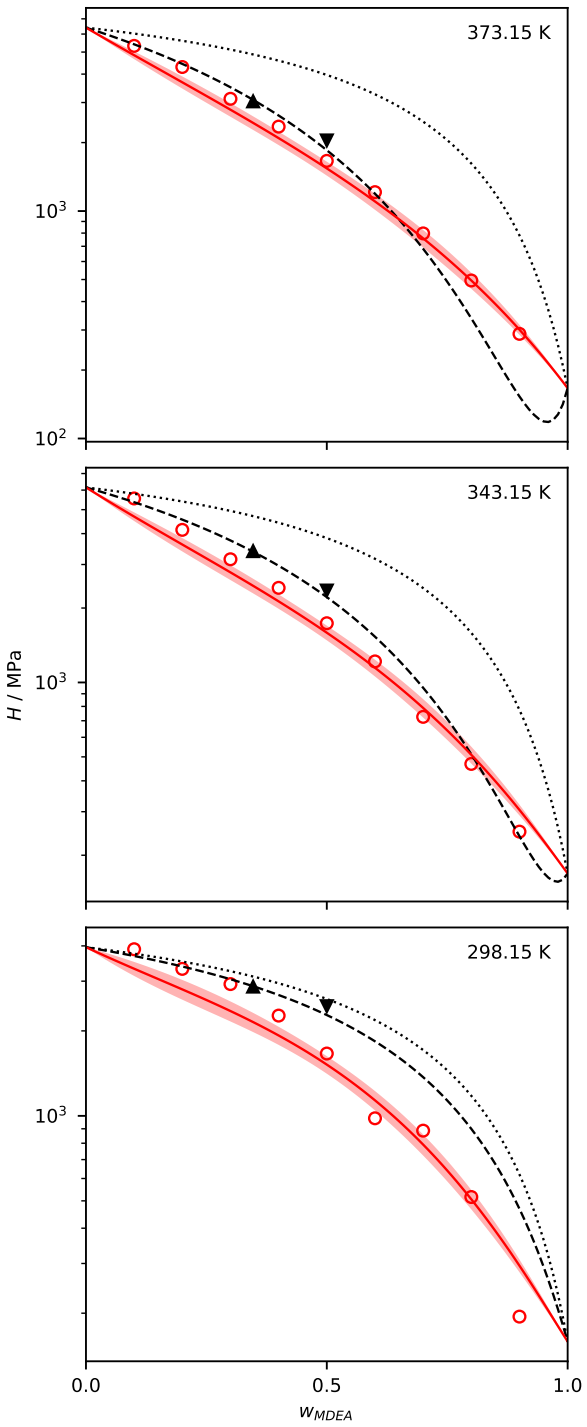

Figure 11: Henry's law constant H as a function of composition for CH_4 in MDEA + water mixtures at 298.15 K, 343.15 K, and 373.15 K. ○: Simulated values, ▲: experimental values from Jou et al. [28], ▼: experimental values from Schmidt et al. [29], ··· ideal mixture, – – correlation of Schmidt et al. [29], — prediction on basis of Eq. 1, using α obtained from molecular simulation and Henry's law constants for pure solvents from Harvey [70] and Jou et al. [27]. The shaded area represents 95% confidence intervals.

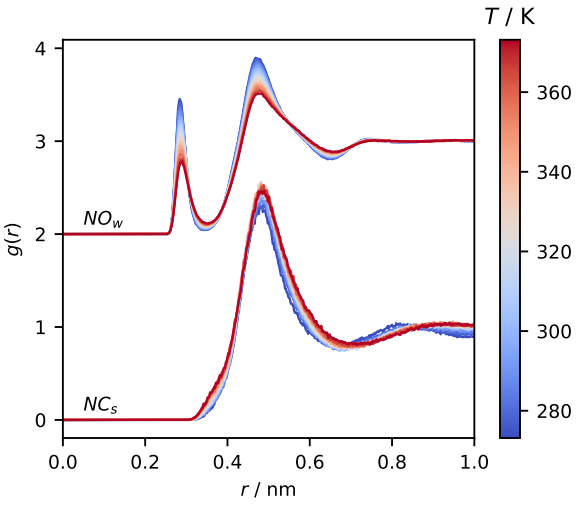

Figure 12: Radial distribution function $g(r)$ as a function of temperature in MDEA + water at 0.5 MDEA weight fraction. The value of $g(r)$ for NO_w is shifted by 2 to reduce overlap.

dependent Margules parameter correlated by Schmidt et al. [29]. While in good agreement with the source data at weight fractions of 0.347 MDEA and 0.5 MDEA, a seemingly spurious minimum in Henry's law constant is predicted at high temperatures. This minimum originates from the peak in α observed in the right panel of Fig. 10 at 0.87 MDEA weight fraction. While the simple expression in Eq. 5 may be sufficient to represent data within a narrow range of concentrations, it is poorly suited to describe solubility in mixed solvents of arbitrary composition.

We again look for the origin of non-ideality in mixture solubility by considering the temperature evolution of radial distribution functions for MDEA + water mixtures. The structure of aqueous alkanolamines has previously been investigated by Orozco et al. [76] near room temperature, finding that alkanolamines preferentially associate with water, rather than other alkanolamines. Hence, we consider radial distribution functions for the atom pair NO_w, with N corresponding to

the tertiary amine in MDEA. In similar fashion as for ethanol + water, we also consider solute-solvent association between CH_4 and MDEA by calculating $g(r)$ for the atom pair NC_s as a function of temperature.

In Fig. 12, radial distribution functions for the atom pairs N_eO_w and N_eC_s are shown for MDEA + water mixtures at 0.5 MDEA weight fraction across the temperature range 273.15 K - 373.15 K. In the case of N_eO_w, there is a primary peak near 0.3 nm corresponding to hydrogen bonding of a single water molecule with the MDEA tertiary amine. The broader secondary peak around 0.5 nm corresponds to solvation of alkane tails and associated alcohol groups. At elevated temperatures, association of MDEA with water decreases, indicating increasing segregation of the two solvents. Conversely, association between MDEA and CH_4 increases with temperature, evidenced by the rising peak in $g(r)$ for NC_s near 0.5 nm. We note that the peak in NC_s is at nearly identical distance as the secondary peak in N_eO_w, suggesting CH_4 is increasingly able to displace H_2O near MDEA at elevated temperatures. For both ethanol + water and MDEA + water, the enhanced solubility of gases at elevated temperatures appears to be driven by increased clustering of organic species.

4. Conclusions

In this article, we have presented a method to obtain the solubility of gases in mixed solvents. The method combines experimental data for solubility in pure solvents with a non-ideal contribution to solubility estimated from molecular simulation. We invoke established theoretical descriptions for gas solubility in non-ideal mixed solvents to explain why such a combination is possible and justified. The resulting non-ideal contribution is primarily a function of solvent-solvent interactions, with only weak dependence on the solute model used. This eliminates the need to optimise solute-solvent interactions, with the same solute model becoming transferable between solvents of any composition.

We test the hybrid method for solubility prediction in two highly non-ideal mixed solvents. Henry's law constants are obtained for CO_2 in ethanol + water mixtures, and for CH_4 in MDEA + water mixtures, in both cases over a wide range of compositions and temperatures. Comparison against experimental data over a wide range of compositions and temperatures gives mean absolute deviations of 6.9% for CO_2 and 27% for CH_4, in each case correctly identifying positive or negative deviations from ideal mixture solubility.

In principle, the method does not require Henry's law constants for pure solvents to be obtained from experimental data. Any reliable source can be used, including thermodynamic models and direct molecular simulation, provided there is sufficient confidence in the model predictions. In all cases, care should be taken to ensure properly converged simulations, as mixture predictions rely on the accuracy of both mixed and pure solvent simulations.

We see several strengths of the present approach for solubility prediction. It can be applied to solvent mixtures of arbitrary composition, implicitly accounting for binary, ternary, and higher-order interactions without the need for complex thermodynamic constructs. It relies on easily measured, often readily accessible experimental data. It eliminates the task of solute model optimisation, greatly simplifying mixed solvent solubility prediction through molecular simulation. These strengths will be advantageous in application to rational design and optimisation of bespoke mixed solvents,

where non-ideal effects can be exploited to improve performance for industrial processes. Through automation of mixture simulations, a large number of compositions can rapidly be screened and combined with, for example, machine learning optimisation strategies. Application of the present method to solvent optimisation will be explored in future studies.

5. Acknowledgements

This work was supported by the European Union's Horizon 2020 Research and Innovation program [Grant Agreement no. 727734]; and the UK Engineering and Physical Sciences Research Council (EPSRC) Doctoral Training Partnership [Grant Number EP/W509644/1]. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (http://www.ecdf.ed.ac.uk/).

Supporting Information

The following Supporting Information is available in the online version of this article: solvent numbers and temperatures used for simulations; tabulated values of ρ, μ_s^E, and excess mixture enthalpy for all simulated solvent compositions and temperatures; examples of the cubic bivariate B-spline method and bootstrap resampling technique used to fit simulation data and generate maximum likelihood estimates; and force field details for gas solutes together with GROMACS formatted force field files for all species considered.

Nomenclature

H	Henry's law constant, MPa
x	Mole fraction
α	Non-ideality parameter (Eq. 1)
R	Molar gas constant, m^3 MPa K^{-1} mol^{-1}
T	Temperature, K
ρ_m	Molar density, mol m^{-3}
ρ	Mass density, kg m^{-3}
M	Molar mass, kg mol^{-1}
μ^E	Excess chemical potential, kJ mol^{-1}
G^E	Excess Gibbs energy, kJ mol^{-1}
α_{ij}	Solvent-solvent Margules parameter
α_{si}	Solute-solvent Margules parameter
$U(\mathbf{r})$	Potential energy, kJ mol^{-1}
ΔU	Potential energy difference, kJ mol^{-1}
$\Delta_{m\to n}G$	Gibbs free energy difference, kJ mol^{-1}
$\langle\cdots\rangle$	Ensemble average
η	Expanded ensemble weight, kJ mol^{-1}
w	Weight fraction
$g(r)$	Radial distribution function

References

[1] T. Rufford, S. Smart, G. Watson, B. Graham, J. Boxall, J. Diniz da Costa, E. May, The removal of CO2 and N2 from natural gas: A review of conventional and emerging process technologies, Journal of Petroleum Science and Engineering 94-95 (2012) 123–154. doi:10.1016/J.PETROL.2012.06.016.

[2] W. M. Bandara, H. Satoh, M. Sasakawa, Y. Nakahara, M. Takahashi, S. Okabe, Removal of residual dissolved methane gas in an upflow anaerobic sludge blanket reactor treating low-strength wastewater at low temperature with degassing membrane, Water Research 45 (11) (2011) 3533–3540. doi:10.1016/J.WATRES.2011.04.030.

[3] G. Brunner, Supercritical fluids: technology and application to food processing, Journal of Food Engineering 67 (1-2) (2005) 21–33. doi:10.1016/J.JFOODENG.2004.05.060.

[4] M. K. Mondal, H. K. Balsora, P. Varshney, Progress and trends in CO2 capture/separation technologies: A review, Energy 46 (1) (2012) 431–441. doi:10.1016/J.ENERGY.2012.08.006.

[5] R. Sander, Compilation of Henry's law constants (version 4.0) for water as solvent, Atmospheric Chemistry and Physics 15 (8) (2015) 4399–4981. doi:10.5194/acp-15-4399-2015.

[6] G. Guevara-Carrion, J. Vrabec, H. Hasse, Prediction of self-diffusion coefficient and shear viscosity of water and its binary mixtures with methanol and ethanol by molecular simulation, The Journal of Chemical Physics 134 (7) (2011) 074508. doi:10.1063/1.3515262.

[7] A. Ghoufi, F. Artzner, P. Malfreyt, Physical Properties and Hydrogen-Bonding Network of Water-Ethanol Mixtures from Molecular Dynamics Simulations, The Journal of Physical Chemistry B 120 (4) (2016) 793–802. doi:10.1021/acs.jpcb.5b11776.

[8] S. M. Melnikov, M. Stein, Molecular Dynamics Study of the Solution Structure, Clustering, and Diffusion of Four Aqueous Alkanolamines, The Journal of Physical Chemistry B 122 (10) (2018) 2769–2778. doi:10.1021/acs.jpcb.7b10322.

[9] J. J. Potoff, J. I. Siepmann, Vapor-liquid equilibria of mixtures containing alkanes, carbon dioxide, and nitrogen, AIChE Journal 47 (7) (2001) 1676–1682. doi:10.1002/aic.690470719.

[10] B. Chen, J. I. Siepmann, Transferable Potentials for Phase Equilibria. 3. Explicit-Hydrogen Description of Normal Alkanes, Journal of Physical Chemistry B 103 (25) (1999) 5370–5379. doi:10.1021/JP990822M.

[11] L. Zhang, J. I. Siepmann, Direct calculation of Henry's law constants from Gibbs ensemble Monte Carlo simulations: Nitrogen, oxygen, carbon dioxide and methane in ethanol, Theoretical Chemistry Accounts 115 (5) (2006) 391–397. doi:10.1007/s00214-005-0073-1.

[12] G. A. Orozco, V. Lachet, A. D. Mackie, Physical Absorption of Green House Gases in Amines: The Influence of Functionality, Structure, and Cross-Interactions, The Journal of Physical Chemistry B 120 (51) (2016) 13136–13143. doi:10.1021/acs.jpcb.6b09819.

[13] G. A. Orozco, V. Lachet, C. Nieto-Draghi, A. D. Mackie, A Transferable Force Field for Primary, Secondary, and Tertiary Alkanolamines, Journal of Chemical Theory and Computation 9 (4) (2013) 2097–2103. doi:10.1021/ct301098s.

[14] H. Docherty, A. Galindo, C. Vega, E. Sanz, A potential model for methane in water describing correctly the solubility of the gas and the properties of the methane hydrate, The Journal of Chemical Physics 125 (7) (2006) 074510. doi:10.1063/1.2335450.

[15] J. L. F. Abascal, C. Vega, A general purpose model for the condensed phases of water: TIP4P/2005, The Journal of Chemical Physics 123 (23) (2005) 234505. doi:10.1063/1.2121687.

[16] H. S. Ashbaugh, N. J. Collett, H. W. Hatch, J. A. Staton, Assessing the thermodynamic signatures of hydrophobic hydration for several common water models, The Journal of Chemical Physics 132 (12) (2010) 124504. doi:10.1063/1.3366718.

[17] J. P. O'Connell, J. M. Prausnitz, Thermodynamics of Gas Solubility in Mixed Solvents, Industrial & Engineering Chemistry Fundamentals 3 (4) (1964) 347–351. doi:10.1021/i160012a012.

[18] K. Shing, K. Gubbins, K. Lucas, Henry constants in non-ideal fluid mixtures, Molecular Physics 65 (5) (1988) 1235–1252. doi:10.1080/00268978800101731.

[19] R. A. Pierotti, A scaled particle theory of aqueous and nonaqueous solutions, Chemical Reviews 76 (6) (1976) 717–726. doi:10.1021/cr60304a002.

[20] A. Gaulhofer, B. Kolbe, J. Gmehling, Thermodynamic properties of ethanol and water III. Description of the different excess functions (gE, hE, cEP) using an empirical model, Fluid Phase Equilibria 39 (2) (1988) 193–209. doi:10.1016/0378-3812(88)85004-0.

[21] A. Idrissi, P. Jedlovszky, Thermodynamics of Mixing Primary Alkanolamines with Water, The Journal of Physical Chemistry B 122 (23) (2018) 6251–6259. doi:10.1021/acs.jpcb.8b01052.

[22] J. K. Kung, F. N. Mazario, J. Joffe, D. Tassios, Prediction of Henry's Constants in Mixed Solvents from Binary Data, Industrial and Engineering Chemistry Process Design and Development 23 (1) (1984) 170–175. doi:10.1021/i200024a028.

[23] G. M. Wilson, Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing, Journal of the American Chemical Society 86 (2) (1964) 127–130. doi:10.1021/ja01056a002.

[24] I. Shulgin, E. Ruckenstein, Henry's constant in mixed solvents from binary data, Industrial and Engineering Chemistry Research 41 (6) (2002) 1689–1694. doi:10.1021/ie010911x.

[25] I. Dalmolin, E. Skovroinski, A. Biasi, M. L. Corazza, C. Dariva, J. V. Oliveira, Solubility of carbon dioxide in binary and ternary mixtures with ethanol and water, Fluid Phase Equilibria 245 (2) (2006) 193–200. doi:10.1016/j.fluid.2006.04.017.

[26] M. A. Postigo, M. Katz, Solubility and thermodynam-

ics of carbon dioxide in aqueous ethanol solutions, Journal of Solution Chemistry 16 (12) (1987) 1015–1024. doi:10.1007/BF00652585.

[27] F.-Y. Jou, A. E. Mather, Solubility of Methane in Methyldiethanolamine, Journal of Chemical & Engineering Data 51 (4) (2006) 1429–1430. doi:10.1021/JE060118G.

[28] F.-Y. Jou, J. J. Carrol, A. E. Mather, F. D. Otto, Solubility of Methane and Ethane in Aqueous Solutions of Methyldiethanolamine, Journal of Chemical & Engineering Data 43 (5) (1998) 781–784. doi:10.1021/JE980003F.

[29] K. A. G. Schmidt, F. Y. Jou, A. E. Mather, Solubility of methane in an aqueous methyldiethanolamine solution (mass fraction 50 %), Journal of Chemical & Engineering Data 53 (8) (2008) 1725–1727. doi:10.1021/je700734p.

[30] S. Garde, A. E. Garca, L. R. Pratt, G. Hummer, Temperature dependence of the solubility of non-polar gases in water, Biophysical Chemistry 78 (1-2) (1999) 21–32. doi:10.1016/S0301-4622(99)00018-6.

[31] William L. Jorgensen, David S. Maxwell, J. Tirado-Rives, Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids, Journal of the American Chemical Society 118 (45) (1996) 11225–11236. doi:10.1021/JA9621760.

[32] R. C. Rizzo, W. L. Jorgensen, OPLS All-Atom Model for Amines: Resolution of the Amine Hydration Problem, Journal of the American Chemical Society 121 (20) (1999) 4827–4836. doi:10.1021/JA984106U.

[33] C. Caleman, P. J. Van Maaren, M. Hong, J. S. Hub, L. T. Costa, D. Van Der Spoel, Force field benchmark of organic liquids: Density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant, Journal of Chemical Theory and Computation 8 (1) (2012) 61–74. doi:10.1021/ct200731v.

[34] M. G. Martin, J. I. Siepmann, Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes, Journal of Physical Chemistry B 102 (14) (1998) 2569–2577. doi:10.1021/JP972543+.

[35] C. Nieto-Draghi, T. de Bruin, J. Pérez-Pellitero, J. Bonet Avalos, A. D. Mackie, Thermodynamic and transport properties of carbon dioxide from molecular simulation, The Journal of Chemical Physics 126 (6) (2007) 064509. doi:10.1063/1.2434960.

[36] C. Vega, J. L. F. Abascal, Simulating water with rigid non-polarizable models: a general perspective, Physical Chemistry Chemical Physics 13 (44) (2011) 19663. doi:10.1039/c1cp22168j.

[37] S. Miyamoto, P. A. Kollman, Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models, Journal of Computational Chemistry 13 (8) (1992) 952–962. doi:10.1002/jcc.540130805.

[38] B. Smit, D. Frenkel, An explicit expression for finite-size corrections to the chemical potential, Journal of Physics: Condensed Matter 1 (44) (1989) 8659–8665. doi:10.1088/0953-8984/1/44/035.

[39] C. B. Barber, D. P. Dobkin, H. Huhdanpaa, H. Huhdanpaa, The quickhull algorithm for convex hulls, ACM Transactions on Mathematical Software 22 (4) (1996) 469–483. doi:10.1145/235815.235821.

[40] K. Hukushima, K. Nemoto, Exchange Monte Carlo Method and Application to Spin Glass Simulations, Journal of the Physical Society of Japan 65 (6) (1996) 1604–1608. doi:10.1143/JPSJ.65.1604.

[41] T. Okabe, M. Kawata, Y. Okamoto, M. Mikami, Replica-exchange Monte Carlo method for the isobaric-isothermal ensemble, Chemical Physics Letters 335 (5-6) (2001) 435–439. doi:10.1016/S0009-2614(01)00055-0.

[42] N. Rathore, M. Chopra, J. J. de Pablo, Optimal allocation of replicas in parallel tempering simulations, The Journal of Chemical Physics 122 (2) (2005) 024111. doi:10.1063/1.1831273.

[43] D. J. Earl, M. W. Deem, Parallel tempering: Theory, applications, and new perspectives, Physical Chemistry Chemical Physics 7 (23) (2005) 3910. doi:10.1039/b509983h.

[44] L. Verlet, Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules, Physical Review 159 (1) (1967) 98–103. doi:10.1103/PhysRev.159.98.

[45] S. Nosé, A molecular dynamics method for simulations in the canonical ensemble, Molecular Physics 52 (2) (1984) 255–268. doi:10.1080/00268978400101201.

[46] W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, Physical Review A 31 (3) (1985) 1695–1697. doi:10.1103/PhysRevA.31.1695.

[47] G. J. Martyna, M. E. Tuckerman, D. J. Tobias, M. L. Klein, Explicit reversible integrators for extended systems dynamics, Molecular Physics 87 (5) (1996) 1117–1157. doi:10.1080/00268979600100761.

[48] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems, The

Journal of Chemical Physics 98 (12) (1993) 10089–10092. doi:10.1063/1.464397.

[49] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, A smooth particle mesh Ewald method, The Journal of Chemical Physics 103 (19) (1995) 8577–8593. doi:10.1063/1.470117.

[50] M. P. Allen, D. J. Tildesley, Computer simulation of liquids, 2nd Edition, Oxford Unitversity Press, Oxford, 2017.

[51] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindah, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1-2 (2015) 19–25. doi:10.1016/j.softx.2015.06.001.

[52] A. Z. Panagiotopoulos, Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble, Molecular Physics 61 (4) (1987) 813–826. doi:10.1080/00268978700101491.

[53] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, P. N. VorontsovVelyaminov, New approach to Monte Carlo calculation of the free energy: Method of expanded ensembles, The Journal of Chemical Physics 96 (3) (1992) 1776–1783. doi:10.1063/1.462133.

[54] B. Widom, Some Topics in the Theory of Fluids, The Journal of Chemical Physics 39 (11) (1963) 2808–2812. doi:10.1063/1.1734110.

[55] J. I. Siepmann, D. Frenkel, Configurational bias Monte Carlo: a new sampling scheme for flexible chains, Molecular Physics 75 (1) (1992) 59–70. doi:10.1080/00268979200100061.

[56] M. R. Shirts, J. D. Chodera, Statistically optimal analysis of samples from multiple equilibrium states, The Journal of Chemical Physics 129 (12) (2008) 124105. doi:10.1063/1.2978177.

[57] T. T. Pham, M. R. Shirts, Optimal pairwise and non-pairwise alchemical pathways for free energy calculations of molecular transformation in solution phase, The Journal of Chemical Physics 136 (12) (2012) 124120. doi:10.1063/1.3697833.

[58] T. T. Pham, M. R. Shirts, Identifying low variance pathways for free energy calculations of molecular transformations in solution phase, The Journal of Chemical Physics 135 (3) (2011) 034114. doi:10.1063/1.3607597.

[59] A. Pohorille, C. Jarzynski, C. Chipot, Good Practices in Free-Energy Calculations, The Journal of Physical Chemistry B 114 (32) (2010) 10235–10253. doi:10.1021/jp102971x.

[60] A. P. Lyubartsev, A. Laaksonen, P. N. Vorontsov-Velyaminov, Free energy calculations for lennard-jones systems and water using the expanded ensemble method a monte carlo and molecular dynamics simulation study, Molecular Physics 82 (3) (1994) 455–471. doi:10.1080/00268979400100344.

[61] J. D. Chodera, M. R. Shirts, Replica exchange and expanded ensemble simulations as Gibbs sampling: Simple improvements for enhanced mixing, The Journal of Chemical Physics 135 (19) (2011) 194110. doi:10.1063/1.3660669.

[62] B. Efron, Bootstrap Methods: Another Look at the Jackknife, The Annals of Statistics 7 (1) (1979) 1–26. doi:10.1214/aos/1176344552.

[63] W. Wagner, A. Pruß, The IAPWS formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use, Journal of Physical and Chemical Reference Data 31 (2) (2002) 387–535. doi:10.1063/1.1461829.

[64] R. L. Rowley, W. V. Wilding, J. L. Oscarson, Y. Yang, N. A. Zundel, T. E. Daubert, R. P. Danner, DIPPR data compilation of pure chemical properties, Design Institute for Physical Properties, AIChE, New York, 2006.

[65] J. Manuel Bernal-García, M. Ramos-Estrada, G. A. Iglesias-Silva, K. R. Hall, Densities and Excess Molar Volumes of Aqueous Solutions of n-Methyldiethanolamine (MDEA) at Temperatures from (283.15 to 363.15) K, Journal of Chemical & Engineering Data 48 (6) (2003) 1442–1445. doi:10.1021/je030120x.

[66] D. G. Friend, D. J. Frurip, E. W. Lemmon, R. E. Morrison, J. D. Olson, L. C. Wilson, Establishing benchmarks for the Second Industrial Fluids Simulation Challenge, Fluid Phase Equilibria 236 (1-2) (2005) 15–24. doi:10.1016/J.FLUID.2005.04.020.

[67] X. Gui, Z. Tang, W. Fei, Solubility of CO2 in Alcohols, Glycols, Ethers, and Ketones at High Pressures from (288.15 to 318.15) K, Journal of Chemical & Engineering Data 56 (5) (2011) 2420–2429. doi:10.1021/je101344v.

[68] S. N. Joung, C. W. Yoo, H. Y. Shin, S. Y. Kim, K.-P. Yoo, C. S. Lee, W. S. Huh, Measurements and correlation of high-pressure VLE of binary CO2-alcohol systems (methanol, ethanol, 2-methoxyethanol and 2-ethoxyethanol), Fluid Phase Equilibria 185 (1-2) (2001) 219–230. doi:10.1016/S0378-3812(01)00472-1.

[69] K. Suzuki, H. Sue, M. Itou, R. L. Smith, H. Inomata, K. Arai, S. Saito, Isothermal vapor-liquid equilibrium data for binary systems at high pressures: carbon dioxide-methanol, carbon dioxide-ethanol, carbon dioxide-1-propanol, methane-ethanol, methane-1-propanol, ethane-ethanol, and ethane-1-propanol

systems, Journal of Chemical & Engineering Data 35 (1) (1990) 63–66. doi:10.1021/je00059a020.

[70] A. H. Harvey, Semiempirical Correlation for Henry's Constants over Large Temperature Ranges, AIChE Journal 42 (5) (1996) 1491–1494. doi:10.1002/aic.690420531.

[71] R. Fernández-Prini, J. L. Alvarez, A. H. Harvey, Henrys Constants and Vapor-Liquid Distribution Constants for Gaseous Solutes in H2O and D2O at High Temperatures, Journal of Physical and Chemical Reference Data 32 (2) (2003) 903–916. doi:10.1063/1.1564818.

[72] R. Crovetto, R. Fernández-Prini, M. L. Japas, Solubilities of inert gases and methane in H2O and in D2O in the temperature range of 300 to 600 K, The Journal of Chemical Physics 76 (2) (1982) 1077–1086. doi:10.1063/1.443074.

[73] N. Nishi, S. Takahashi, M. Matsumoto, A. Tanaka, K. Muraya, T. Takamuku, T. Yamaguchi, Hydrogen-Bonded Cluster Formation and Hydrophobic Solute Association in Aqueous Solutions of Ethanol, The Journal of Physical Chemistry 99 (1) (1995) 462–468. doi:10.1021/j100001a068.

[74] S. Dixit, J. Crain, W. C. K. Poon, J. L. Finney, A. K. Soper, Molecular segregation observed in a concentrated alcoholwater solution, Nature 416 (6883) (2002) 829–832. doi:10.1038/416829a.

[75] S. Lenton, N. H. Rhys, J. J. Towey, A. K. Soper, L. Dougan, Temperature-Dependent Segregation in Alcohol-Water Binary Mixtures Is Driven by Water Clustering, The Journal of Physical Chemistry B 122 (32) (2018) 7884–7894. doi:10.1021/ACS.JPCB.8B03543.

[76] G. A. Orozco, V. Lachet, A. D. Mackie, A molecular simulation study of aqueous solutions of amines and alkanolamines: mixture properties and structural analysis, Molecular Simulation 40 (1-3) (2014) 123–133. doi:10.1080/08927022.2013.845297.